Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

A fast reduced-rank sound zone control algorithm using the conjugate gradient method

Shi, Liming; Lee, Taewoong; Zhang, Lijun; Nielsen, Jesper Kjær; Christensen, Mads Græsbøll

Published in:
International Conference on Acoustics, Speech, and Signal Processing

DOI (link to publication from Publisher):
10.1109/ICASSP40776.2020.9054461

Publication date:
2020

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):
Shi, L., Lee, T., Zhang, L., Nielsen, J. K., & Christensen, M. G. (2020). A fast reduced-rank sound zone control algorithm using the conjugate gradient method. In *International Conference on Acoustics, Speech, and Signal Processing* (pp. 436-440). [9054461] IEEE. I E E E International Conference on Acoustics, Speech and Signal Processing. Proceedings https://doi.org/10.1109/ICASSP40776.2020.9054461

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

A FAST REDUCED-RANK SOUND ZONE CONTROL ALGORITHM USING THE CONJUGATE GRADIENT METHOD

Liming Shi,[1] *Taewoong Lee,*[1] *Lijun Zhang,*[2] *Jesper Kjær Nielsen,*[1] *Mads Græsbøll Christensen*[1]

[1]Audio Analysis Lab, CREATE, Aalborg University, {ls, tlee, jkn, mgc}@create.aau.dk
[2] School of Marine Science and Technology, Northwestern Polytechnical University, zhanglj7385@nwpu.edu.cn

ABSTRACT

Sound zone control enables different users to enjoy different audio contents in the same acoustic environment. Generalized eigenvalue decomposition (GEVD)-based methods allow us to control the trade-off between the acoustic contrast (AC) and signal distortion (SD). However, such methods have a high computational complexity. In this paper, we propose a fast reduced-rank sound zone control algorithm using the conjugate gradient (CG) method. Instead of using the eigenvectors as the basis for the solution space, the search directions in the CG method are used to reduce the computational complexity. Then, a low dimensional EVD is applied to obtain the sub-optimal control filter coefficients. The dark zone power can be adjusted by a parameter, which implicitly controls the trade-off between the AC and SD. Compared with GEVD-based methods, experimental results show that the proposed algorithm has a degradation of performance (4–5 dB) in terms of AC or SD but a high improvement on computational efficiency.

Index Terms— Computational complexity, conjugate gradient, reduced-rank, sound zone control, variable span trade-off filter

1. INTRODUCTION

After the concept of personal sound was first proposed about two decades ago [1], the creation of sound zones became an active research field [2, 3] and various applications have been studied [4–10]. The main idea behind sound zones is to generate different listening areas in the same acoustic space for different audio contents with minimum or ideally no disturbance between the areas using a loudspeaker array, e.g., a soundbar. In order to achieve this goal, the following two zones for a single audio content are typically considered: a bright zone (or a listening zone) which is a confined area whose acoustic potential energy is maximized or where a desired sound field is reproduced and a dark zone (or a quiet zone) which is another confined area whose acoustic potential energy is minimized. By exploiting the superposition principle, multiple bright zones can be obtained. Broadly, three different control strategies have been extensively studied: acoustic contrast control (ACC), pressure matching (PM), and mode matching. Note that the mode matching(or the modal domain) approach represents the sound fields in zones via a spatial harmonic expansion [11, 12] to find the control filters, but we here focus on the first two methods. ACC maximizes the acoustic contrast which describes the acoustic potential energy ratio between the bright and dark zones [13]. Although ACC guarantees the maximum acoustic contrast at the corresponding frequency, it does not ensure a proper distribution of sound pressure across the zones and the frequency of interest which might degrade the reproduced sound fields. This issue becomes more problematic when it is designed in the time domain since the maximum acoustic contrast is often obtained by filtering out all in the signal except for one or a few sinusoidal components which causes a significant signal distortion [14]. In order to mitigate this issue, various techniques have been studied, e.g., [14–16]. In contrast to ACC, PM minimizes the reproduction error which is the difference between the reproduced and desired sound fields in the bright and dark zones. However, it has a low acoustic contrast. A combination of the ACC and PM methods has been studied in both the frequency domain [17] and the time domain [18] in order to trade-off the signal distortion in the bright zone and the dark zone power, which implicitly changes the acoustic contrast accordingly. Recently, motivated by the variable span linear filters for speech enhancement [19, 20], a framework referred to as variable span trade-off (VAST) based on GEVD for sound zone control has been proposed [21] which allows one to trade-off the signal distortion and acoustic contrast. Moreover, the traditional approaches including ACC, PM, and ACC-PM can be seen as special cases of VAST. However, in VAST, computing GEVD with high data dimension leads to a high computational complexity. Moreover, although VAST is derived based on a constrained optimization problem, the Lagrange multiplier has not yet been computed explicitly, rather it is typically chosen as a fixed value.

In this paper, we propose a fast reduced-rank sound zone control algorithm using the CG method. Using the CG method for obtaining the basis for the solution space, instead of the GEVD, the computational complexity can be reduced. To control the dark zone power, a low dimensional EVD is used to obtain the estimate of the sound zone control filter. For both VAST and the proposed CG-based algorithms, we show that a linear search method can be used to find the Lagrange multiplier for a specifically chosen value of the dark zone power. Both computational complexity analysis and experimental results are used to verify the performance of the proposed algorithm.

2. FUNDAMENTALS

In this section, we first introduce the VAST approach for sound zone control and then briefly describe the CG method for obtaining a basis of the solution space for the proposed low rank method.

2.1. Variable span trade-off filter for sound zone control

We consider the problem of generating a bright zone and a dark zone in an enclosed space with measured/known room impulse responses (RIRs) using L loudspeakers and L finite impulse response (FIR) filters (a.k.a., control filters). The reproduced sound signal at the $m^{\text{th}}, 1 \leq m \leq M^{\text{C}}$ sampling point in one of the sound zones can be

written as

$$y_n^{m,\mathrm{C}} = \sum_{l=1}^{L} x_n * q_n^l * h_n^{m,l,\mathrm{C}}, \tag{1}$$

where superscript $(\cdot)^{\mathrm{C}}, \mathrm{C} \in \{\mathrm{B}, \mathrm{D}\}$ is the zone index, C=B and C=D denote the bright zone and dark zone, respectively, x_n denotes the input sound signal, q_n^l denotes the l^{th} control filter of length J for the l^{th} loudspeaker, $h_n^{m,l,\mathrm{C}}$ denotes the RIR of length K from the l^{th} loudspeaker to the m^{th} sampling point in zone C, $*$ denotes the convolution operator and n denotes the time index. For simplicity, we have left out the B and D superscripts when not absolutely necessary. Collecting N samples and writing (1) in a matrix form, we have

$$\mathbf{y}^m = \mathbf{U}^m \mathbf{q}, \tag{2}$$

where

$$\mathbf{y}^m = [y_N^m, \cdots, y_1^m]^T,\ \mathbf{q} = \left[\mathbf{q}^{1T}, \cdots, \mathbf{q}^{LT}\right]^T \in \mathcal{R}^{LJ\times 1},$$

$$\mathbf{q}^l = \left[q_1^l, \cdots, q_J^l\right]^T, \mathbf{U}^m = \left[\mathbf{U}^{m,1}, \cdots, \mathbf{U}^{m,L}\right] \in \mathcal{R}^{N\times LJ},$$

$$\mathbf{U}^{m,l} = \left[\mathbf{u}_N^{m,l}, \mathbf{u}_{N-1}^{m,l}, \cdots, \mathbf{u}_1^{m,l}\right]^T \in \mathcal{R}^{N\times J},$$

$$\mathbf{u}_n^{m,l} = \left[u_n^{m,l}, \cdots, u_{n-J+1}^{m,l}\right]^T, u_n^{m,l} = x_n * h_n^{m,l}.$$

Assuming the desired sound signal for a sound zone is $\mathbf{d}^m$, where $\mathbf{d}^m$ is defined similarly to $\mathbf{y}_n^m$, using (2), the least squares cost function for this zone can be expressed as

$$J(\mathbf{q}) = \sum_{m=1}^{M} \|\mathbf{d}^m - \mathbf{y}^m\|_2^2 = \mathbf{q}^T\mathbf{R}\mathbf{q} - 2\mathbf{q}^T\mathbf{r} + \kappa, \tag{3}$$

where $\kappa = \sum_{m=1}^{M} \|\mathbf{d}^m\|_2^2$, $\mathbf{R}$ and $\mathbf{r}$ denote the spatial autocorrelation matrix and spatial cross-correlation vector given by $\mathbf{R} = \sum_{m=1}^{M} (\mathbf{U}^m)^T\mathbf{U}^m$ and $\mathbf{r} = \sum_{m=1}^{M} (\mathbf{U}^m)^T\mathbf{d}^m$. Note that, to reduce the signal power in the dark zone, $\mathbf{d}^{m,\mathrm{D}} = \mathbf{0}$ is commonly applied (leading to $\mathbf{r}^{\mathrm{D}} = \mathbf{0}$), and we use this setting throughout this paper [21]. Recently, motivated by a framework proposed in speech enhancement [19, 22], a VAST method for sound zone control has been proposed [21]. The VAST filter can be derived by solving the following optimization problem:

$$\hat{\mathbf{q}} = \arg\min_{\mathbf{q}} J^{\mathrm{B}}(\mathbf{q}) \quad \text{s.t.} \quad J^{\mathrm{D}}(\mathbf{q}) \le \sigma^2, \tag{4}$$

where σ^2 denotes an upper bound for the dark zone power. Plugging (3) into (4), using the method of Lagrange multipliers and following the derivation in [21], the optimal control filter can be written as

$$\hat{\mathbf{q}} = \sum_{i=1}^{V} \frac{\mathbf{u}_i^T\mathbf{r}^{\mathrm{B}}}{\lambda_i + \mu}\mathbf{u}_i, \tag{5}$$

where $\mathbf{u}_i$ and λ_i ($\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{LJ}$) denote the eigenvector and the eigenvalue, respectively, for the generalized eigenvalue problem $\mathbf{R}^{\mathrm{B}}\mathbf{u}_i = \lambda_i\mathbf{R}^{\mathrm{D}}\mathbf{u}_i, 1 \le i \le LJ$ [23], μ denotes the Lagrange multiplier and V denotes the total number of ranks. The Lagrange multiplier μ is constrained by $\hat{\mathbf{q}}^T\mathbf{R}^{\mathrm{D}}\hat{\mathbf{q}} \le \sigma^2$. It is further shown in [20], the parameter V can be used to control the trade-off between the acoustic contrast and signal distortion. More specifically, a larger V leads to a lower signal distortion but a decreased acoustic contrast, and vice versa. In extreme cases, when $V = 1$, the VAST filter reduces to the ACC approach proposed in [24] and when $V = LJ$, the VAST filter reduces to the PM approach proposed in [25]. The main disadvantage of VAST is its computational complexity. The plain implementation of the GEVD requires $\mathcal{O}(L^3J^3)$ operations. In this paper, we propose an alternative low rank sound zone control method using the CG method with $\mathcal{O}(VL^2J^2)$ operations. We first introduce the classical CG method in the next subsection.

Algorithm 1 The conjugate gradient method

Initiate $\mathbf{q}_1 = \mathbf{0}, \mathbf{d}_1 = \mathbf{r}_1 = \mathbf{r}, g_1 = \mathbf{r}_1^T\mathbf{r}_1$
for $p = 1, 2, \cdots, P$ **do**
  $\mathbf{c}_p = \mathbf{R}\mathbf{d}_p$ — $\mathcal{O}(P^2)$
  $\alpha_p = \frac{g_p}{\mathbf{d}_p^T\mathbf{c}_p}$ — $\mathcal{O}(P)$
  $\mathbf{q}_{p+1} = \mathbf{q}_p + \alpha_p\mathbf{d}_p$ — $\mathcal{O}(P)$
  $\mathbf{r}_{p+1} = \mathbf{r}_p - \alpha_p\mathbf{c}_p$ — $\mathcal{O}(P)$
  $g_{p+1} = \mathbf{r}_{p+1}^T\mathbf{r}_{p+1}$ — $\mathcal{O}(P)$
  $\beta_{p+1} = \frac{g_{p+1}}{g_p}$ — $\mathcal{O}(1)$
  $\mathbf{d}_{p+1} = \mathbf{r}_{p+1} + \beta_{p+1}\mathbf{d}_p$ — $\mathcal{O}(P)$
end for

2.2. The conjugate gradient (CG) method

Considering solving a linear equation $\mathbf{R}\mathbf{q} = \mathbf{r}$, where $\mathbf{R}$ is a $P \times P$ full-rank symmetric matrix and $P = LJ$, the plain implementation of $\mathbf{q} = \mathbf{R}^{-1}\mathbf{r}$ requires $\mathcal{O}(P^3)$ operations. It can be easily shown that solving $\mathbf{R}\mathbf{q} = \mathbf{r}$ is equivalent to finding the optimal point for $\arg\min_{\mathbf{q}} f(\mathbf{q}) = \mathbf{q}^T\mathbf{R}\mathbf{q} - 2\mathbf{r}^T\mathbf{q}$ [26]. The CG method iteratively minimizes the quadratic cost function $f(\mathbf{q})$ along a set of $\mathbf{R}$-orthogonal search directions. The CG algorithm and its computation complexity for each step is shown in Algorithm 1, where $\mathbf{d}_n$ and $\mathbf{r}_n$ denote the search direction and residual vector at the n^{th} iteration, respectively. One important property of the CG algorithm is that the search directions $\mathbf{d}_n, n \ge 1$ are $\mathbf{R}$-orthogonal, i.e.,

$$\mathbf{D}_p^T\mathbf{R}\mathbf{D}_p = \mathbf{\Lambda}_p^{\mathrm{CG}}, \quad 1 \le p \le P, \tag{6}$$

where $\mathbf{\Lambda}_p^{\mathrm{CG}} = \mathrm{diag}\{\mathbf{d}_1^T\mathbf{A}\mathbf{d}_1, \cdots, \mathbf{d}_p^T\mathbf{A}\mathbf{d}_p\}$, and $\mathbf{D}_p = [\mathbf{d}_1, \cdots, \mathbf{d}_p]$. Using (6) and setting $p = P$, we can easily obtain

$$\mathbf{R}^{-1}\mathbf{r} = \mathbf{D}_P(\mathbf{\Lambda}_{\mathrm{P}}^{\mathrm{CG}})^{-1}\mathbf{D}_P^T\mathbf{r}. \tag{7}$$

Also, as can be seen from Algorithm 1, by setting $\mathbf{q}_1 = \mathbf{0}$ and using the iterative equation $\mathbf{q}_{p+1} = \mathbf{q}_p + \alpha_p\mathbf{d}_p$, we have

$$\mathbf{R}^{-1}\mathbf{r} = \mathbf{D}_P\boldsymbol{\alpha}_P, \tag{8}$$

where $\boldsymbol{\alpha}_P = [\alpha_1, \cdots, \alpha_P]^T$ denotes the weighting vector for search directions. Comparing (7) and (8), we yield

$$\boldsymbol{\alpha}_P = (\mathbf{\Lambda}_P^{\mathrm{CG}})^{-1}\mathbf{D}_P^T\mathbf{r}. \tag{9}$$

3. REDUCED-RANK SOUND ZONE CONTROL USING THE CG METHOD

In this paper, we propose a fast reduced-rank (e.g., V-rank and $V \le LJ$) sound zone control algorithm using the CG method and only a $V-$dimensional EVD is required for the proposed method. Substituting (3) into (4), the optimization problem becomes

$$\hat{\mathbf{q}} = \arg\min_{\mathbf{q}} \mathbf{q}^T\mathbf{R}^{\mathrm{B}}\mathbf{q} - 2\mathbf{q}^T\mathbf{r}^{\mathrm{B}} + \kappa^{\mathrm{B}} \quad \text{s.t.} \quad \mathbf{q}^T\mathbf{R}^{\mathrm{D}}\mathbf{q} \le \sigma^2. \tag{10}$$

We constrain the solution space for $\mathbf{q}$ to a V-dimensional subspace $\mathcal{K}_V$ with a basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_V\}$. Denote $\mathbf{U}_V = [\mathbf{u}_1, \cdots, \mathbf{u}_V]$ and let $\mathbf{q} = \mathbf{U}_V\mathbf{z}$. Then, the constrained optimization problem (10) is equivalent to

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \mathbf{z}^T(\mathbf{U}_V{}^T\mathbf{R}^{\mathrm{B}}\mathbf{U}_V)\mathbf{z} - 2\mathbf{z}^T(\mathbf{U}_V{}^T\mathbf{r}^{\mathrm{B}}), \tag{11}$$
$$\text{s.t.} \quad \mathbf{z}^T(\mathbf{U}_V{}^T\mathbf{R}^{\mathrm{D}}\mathbf{U}_V)\mathbf{z} \le \sigma^2.$$

The basis matrix $\mathbf{U}_V$ can be chosen in different ways. For example, if we set $V = LJ$ and $\mathbf{U}_V = \mathbf{I}$, the solution to (11) is simply

the PM solution [25]. Another example is to set each column vector of $\mathbf{U}_V$ to the eigenvector for the generalized eigenvalue problem $\mathbf{R}^{\mathrm{B}}\mathbf{u}_i = \lambda_i \mathbf{R}^{\mathrm{D}}\mathbf{u}_i$, then we can obtain the VAST filter. In this paper, we use the search directions in the CG method to form the basis matrix, i.e., $\mathbf{U}_V = [\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_V]$, where $\mathbf{d}_v, 1 \leq v \leq V$ denotes the search directions in Algorithm 1 for solving the problem $\mathbf{R}^{\mathrm{B}}\mathbf{x} = \mathbf{r}^{\mathrm{B}}$. There are multiple reasons for this choice. First, a filter with the smallest signal distortion for the bright zone can be obtained by solving $\mathbf{R}^{\mathrm{B}}\mathbf{q} = \mathbf{r}^{\mathrm{B}}$. Secondly, the search directions are $\mathbf{R}^{\mathrm{B}}$-orthogonal, making the optimization problem (11) simpler. Thirdly, obtaining a V-dimensional basis matrix $\mathbf{U}_V$ by using the CG method ($\mathcal{O}(VL^2J^2 + 6VLJ)$) is computationally simpler than the GEVD ($\mathcal{O}(L^3J^3)$) method. Using (6), the spatial auto-correlation matrix of the bright zone can be diagonalized to

$$\mathbf{U}_V{}^T \mathbf{R}^{\mathrm{B}} \mathbf{U}_V = \mathbf{\Lambda}_V^{\mathrm{CG}}. \tag{12}$$

Moreover, using (9), we can obtain

$$\boldsymbol{\alpha}_V = (\mathbf{\Lambda}_V^{\mathrm{CG}})^{-1} \mathbf{U}_V^T \mathbf{r}^{\mathrm{B}}, \tag{13}$$

where $\boldsymbol{\alpha}_V = [\alpha_1, \cdots, \alpha_V]^T$. Substituting (12) and (13) into (11) and letting $\mathbf{t} = (\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\mathbf{z}$, we obtain

$$\begin{aligned} \hat{\mathbf{t}} &= \arg\min_{\mathbf{t}} \mathbf{t}^T\mathbf{t} - 2\mathbf{t}^T((\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\boldsymbol{\alpha}_V), \\ \text{s.t.} \quad & \mathbf{t}^T\mathbf{M}_V\mathbf{t} \leq \sigma^2, \end{aligned} \tag{14}$$

where

$$\mathbf{M}_V = (\mathbf{\Lambda}_V^{\mathrm{CG}})^{-1/2} \mathbf{U}_V{}^T \mathbf{R}^{\mathrm{D}} \mathbf{U}_V (\mathbf{\Lambda}_V^{\mathrm{CG}})^{-1/2}. \tag{15}$$

By using the method of Lagrange multipliers, a stationary point of the constrained optimization problem (14) can be obtained as

$$\hat{\mathbf{t}} = (\mathbf{I} + \mu\mathbf{M}_V)^{-1}((\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\boldsymbol{\alpha}_V), \tag{16}$$

where μ denotes the Lagrange multiplier, and it needs to satisfy the constraint $\hat{\mathbf{t}}^T\mathbf{M}_V\hat{\mathbf{t}} \leq \sigma^2$. Using a V-dimensional eigenvalue decomposition (EVD) to diagonalize $\mathbf{M}_V \in \mathcal{R}^{V\times V}$, i.e.,

$$\mathbf{F}_V^T \mathbf{M}_V \mathbf{F}_V = \mathbf{\Lambda}_V^{\mathrm{Eig}}, \tag{17}$$

where the column vectors of $\mathbf{F}_V$ are orthonormal eigenvectors, and the diagonal elements of $\mathbf{\Lambda}_V^{\mathrm{Eig}}$ contain the corresponding eigenvalues. Substituting (17) into (16), we then obtain

$$\hat{\mathbf{t}} = \mathbf{F}_V(\mathbf{I} + \mu\mathbf{\Lambda}_V^{\mathrm{Eig}})^{-1}\mathbf{F}_V^T((\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\boldsymbol{\alpha}_V), \tag{18}$$

with the constraint

$$f(\mu) = \sum_{v=1}^{V} \frac{\lambda_v^{\mathrm{Eig}} c_v^2}{(1+\mu\lambda_v^{\mathrm{Eig}})^2} \leq \sigma^2, \tag{19}$$

where c_v denotes the v^{th} element of $\mathbf{F}_V^T((\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\boldsymbol{\alpha}_V)$. Because $\mathbf{M}_V$ is a nonnegative definite matrix, the eigenvalue λ_v^{Eig} is nonnegative. Therefore, $f(\mu)$ is monotonically decreasing for $\mu \geq 0$ and the solution for $f(\mu) \leq \sigma^2$ exists since $\lim_{\mu\to\infty} f(\mu) = 0$. In this paper, the Newton's method [27] is applied to find the solution for the Lagrange multiplier μ. By using $\mathbf{q} = \mathbf{U}_V\mathbf{z}$ and $\mathbf{t} = (\mathbf{\Lambda}_V^{\mathrm{CG}})^{1/2}\mathbf{z}_n$, the estimated control filter can be expressed as

$$\hat{\mathbf{q}} = \mathbf{U}_V(\mathbf{\Lambda}_V^{\mathrm{CG}})^{-1/2}\hat{\mathbf{t}}. \tag{20}$$

The proposed Reduced-Rank sound zone control algorithm using the CG method (RR-CG) is summarized in Algorithm 2, where n_{iter} denotes the number of iterations for the Newton's method.

In VAST [21], the Lagrange multiplier μ is set to a fixed value (e.g., $\mu = 0.8$). In fact, when the basis matrix $\mathbf{U}_V$ is formed by the eigenvectors using GEVD (denoted as $\mathbf{U}_{V,\mathrm{GEVD}}$), the matrices

Algorithm 2 The proposed Reduced-Rank sound zone control algorithm using the CG method (RR-CG)

1: Initiate the number of ranks V and σ^2.
2: Run the CG method V iterations for the problem $\mathbf{R}^{\mathrm{B}}\mathbf{x} = \mathbf{r}^{\mathrm{B}}$, and store $\mathbf{U}_V$, $\boldsymbol{\alpha}_V$, and $\mathbf{\Lambda}_V^{\mathrm{CG}}$. $\mathcal{O}(V(LJ)^2)$
3: Form $\mathbf{M}_V$ based on (15). $\mathcal{O}(V(LJ)^2)$
4: Compute the EVD of $\mathbf{M}_V$. $\mathcal{O}(V^3)$
5: Use the Newton's method to find μ based on (19). $\mathcal{O}(n_{\text{iter}}V)$
6: Compute $\hat{\mathbf{t}}$ based on (18). $\mathcal{O}(V^2)$
7: Compute $\hat{\mathbf{q}}$ based on (20). $\mathcal{O}(VLJ)$

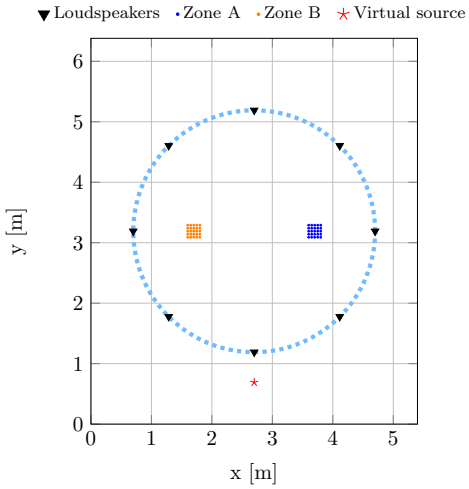

Fig. 1. An example of sound zone control system setup with eight loudspeakers, 25 control points in each zone and two virtual sources on the same location.

$\mathbf{R}^{\mathrm{B}}$ and $\mathbf{R}^{\mathrm{D}}$ can be jointly diagonalized by $\mathbf{U}_{V,\mathrm{GEVD}}$. Using this property and solving (11), the optimal μ can be expressed as

$$g(\mu) = \sum_{v=1}^{V} \frac{c_{v,\mathrm{GEVD}}^2}{(\mu+\lambda_v)^2} \leq \sigma^2, \tag{21}$$

where $c_{v,\mathrm{GEVD}}$ denotes the v^{th} element of $\mathbf{U}_{V,\mathrm{GEVD}}^T\mathbf{r}^{\mathrm{B}}$, and λ_v denotes the v^{th} eigenvalue. Again, the Newton's search can be used to find the solution for (21). We refer to this algorithm as VAST with Optimal μ (VAST-O). The advantage of using the tuning parameter σ^2 in VAST-O, instead of the μ in VAST, is that σ^2 has physical meaning and it can be seen as the upper bound for the reconstructed dark zone power.

4. RESULTS

In this section, the performance of the RR-CG and VAST-O methods [1] for sound zone control is compared with the ACC, PM, and VAST methods [21] on real speech signals. Three types of performance measures are used to quantify the experimental results, i.e., acoustic contrast (AC), signal distortion (SD) and target to interferer ratio (TIR) (see [28] for more on this). As shown in Fig. 1, we consider a system which consists of a circular array with eight evenly distributed loudspeakers, two zones, and a virtual source. Each of the zones are spatially sampled by a 2-D square grid of 5×5 control points spaced by 5 cm is used. We assume that all loudspeakers

[1] MATLAB code in `https://github.com/LimingShi/RR-CG`.

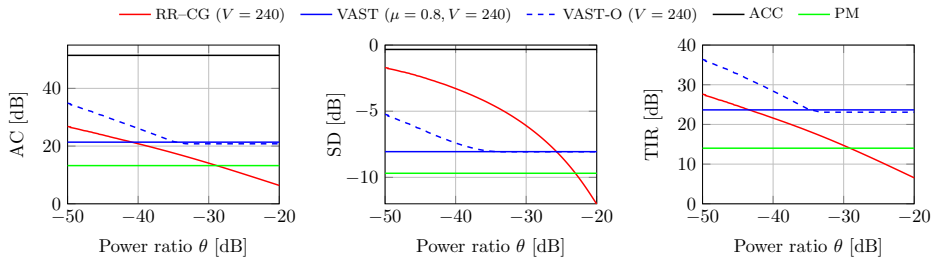

Fig. 2. The AC, SD and TIR performances of the proposed RR-CG for different choices of σ^2 by adjusting the power ratio θ.

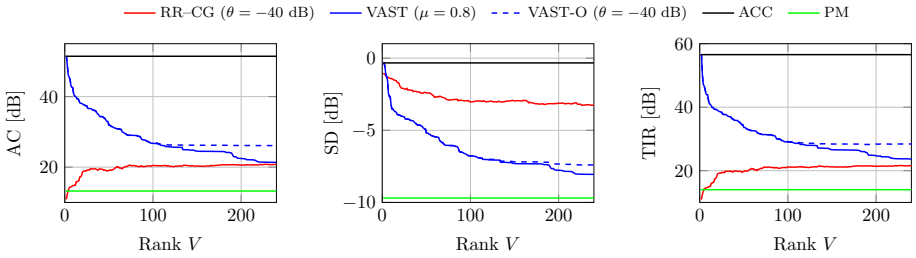

Fig. 3. The AC, SD and TIR performances of the proposed RR-CG for different choices of σ^2 by adjusting the rank V ($LJ = 960$).

and control points are located on the same plane. For simplicity, we assume that no reverberation is present, i.e., free field. The lengths of the control filters and impulse responses are set to $J = 120$ and $K = 1600$, respectively. The sampling frequency is set to 8 kHz. Two six-seconds long speech signals from the movie "Zootopia" in two different languages, i.e., English and Danish, are used as the input signal x_n. The signal powers for the two input signals are set to be the same. In RR-CG and VAST-O, the maximum number of iterations for the Newton's method is set to 10^4, the in initial value for μ is set to 0 and the stopping threshold is set to 10^{-7}.

First, the performance of the RR-CG and VAST-O is tested for different choices of the upper bound for the dark zone power σ^2, which is computed based on $\sigma^2 = \theta(\mathbf{q}^u)^T\mathbf{R}^{\mathrm{D}}\mathbf{q}^u$, where $\mathbf{q}^u = \mathbf{1}_L \otimes \mathbf{i}_1$ denotes the uncontrolled filter, $\mathbf{1}_L$ denotes the all-ones vector with L elements, $\mathbf{i}_1$ denotes the all-zeros vector except the first element is one, and θ denotes the power ratio controlling the amount of reduction for the dark zone power. The rank V for the RR-CG, VAST-O and VAST is set to $LJ/4$ (i.e., 240). The Lagrange multiplier for VAST is set to $\mu = 0.8$. The AC, SD, and TIR for the proposed RR-CG, VAST-O, VAST, ACC, and PM are shown in Fig. 2. As can be seen, the ACC approach has the largest AC and TIR, but the highest SD. The PM has a low SD, but small AC and TIR. By using one forth of the eigenvectors, the AC, TIR, and SD of the VAST approach is between ACC and PM. With an increasing θ, the AC, TIR and SD of the VAST-O decreases until θ is larger than a threshold, from where the optimal μ is computed as 0 (i.e., the dark zone power constraint is satisfied using the chosen V). For the proposed RR-CG, with an increasing θ, both the AC and TIR decrease but the SD becomes lower. When the AC for the RR-CG and VAST is set to be the same, i.e,. using $\theta \approx 10^{-4}$, the SD performance of the RR-CG is 5 dB larger than VAST. When the AC for the RR-CG and PM is set to be the same, i.e,. using $\theta \approx 5 \times 10^{-3}$, the SD performance of the RR-CG is 4.5 dB higher than PM.

Secondly, the performance of RR-CG, VAST-O and VAST is tested for different choices of the rank V with $\theta = -40$ dB for RR-CG and VAST-O. The experimental results are shown in Fig. 3. For the proposed RR-CG, the AC and TIR increase, whereas the SD decreases with an increasing V. When $V \approx 50$ ($LJ = 960$), RR-CG starts to converge.

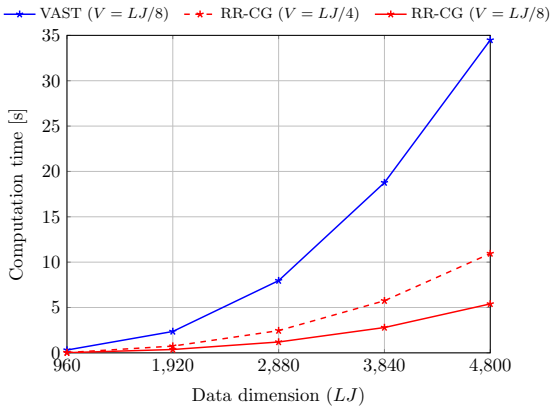

Fig. 4. The computation time for VAST and the proposed RR-CG using the same rank V.

Thirdly, we verify the performance of the proposed RR-CG method in terms of processing time. All timings are computed on a 3.6 GHz Intel(R) Core(TM) i7-4790 CPU with Ubuntu Linux Kernel 4.4.0-97-generic and MATLAB R2017b. The computation time is evaluated as averaged results for 100 Monte Carlo trials with different data dimensions (LJ). The results are shown in Fig. 4. As can be seen, the proposed RR-CG is close to seven times faster than VAST when $V = LJ/8$.

5. CONCLUSION

A fast reduced-rank sound zone control approach using the CG method is proposed. The CG method is applied to form the basis for the solution space due to its low computational complexity compared with the GEVD. For both the CG and GEVD-based approaches, we present a method to control the trade-off between the acoustic contrast and signal distortion by adjusting the upper bound for the reconstructed dark zone power. The proposed CG-based approach has around 4–5 dB lower performance than traditional GEVD-based sound zone control approaches in terms of acoustic contrast and signal distortion, but it features a low computational complexity.

6. REFERENCES

[1] W. F. Druyvesteyn and J. Garas, "Personal sound," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 685–701, Sep. 1997.

[2] T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, "Personal sound zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 81–91, Mar. 2015.

[3] W. Zhang, P. Samarasinghe, H. Chen, and T. D. Abhayapala, "Surround by sound: A review of spatial audio recording and reproduction," *Appl. Sci.*, vol. 7, no. 6, p. 532, May 2017.

[4] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2091–2097, Apr. 2009.

[5] J.-M. Lee, T. Lee, J.-Y. Park, and Y.-H. Kim, "Generation of a private listening zone; acoustic parasol," in *20th Int. Congr. Acoust.*, Sydney, Australia, Aug. 2010.

[6] J. Cheer, S. J. Elliott, Y. Kim, and J.-W. Choi, "Practical implementation of personal audio in a mobile device," *J. Audio Eng. Soc.*, vol. 61, no. 5, pp. 290–300, Jun. 2013.

[7] J. Cheer and S. J. Elliott, "Design and implementation of a personal audio system in a car cabin," in *21st Int. Congr. Acoust.*, Montreal, QC, Canada, Aug. 2013.

[8] X. Liao, J. Cheer, S. J. Elliott, and S. Zheng, "Design array of loudspeakers for personal audio system in a car cabin," in *Proc. 23rd Int. Congr. Sound Vib.*, Athens, Greece, Jul. 2016.

[9] H. So and J.-W. Choi, "Subband optimization and filtering technique for practical personal audio systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Brighton, UK, May 2019, pp. 8494–8498.

[10] J.-H. Chang and W.-H. Cho, "Evaluation of independent sound zones in a car," in *Proceedings of the 23rd International Congress on Acoustics*, Aachen, Germany, Sep. 2019.

[11] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, Aug. 2011.

[12] W. Zhang, T. D. Abhayapala, T. Betlehem, and F. M. Fazi, "Analysis and control of multi-zone sound field reproduction using modal-domain approach," *J. Acoust. Soc. Am.*, vol. 140, no. 3, pp. 2134–2144, Sep. 2016.

[13] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, Apr. 2002.

[14] Y. Cai, M. Wu, and J. Yang, "Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2013, pp. 341–345.

[15] Y. Cai, M. Wu, L. Liu, and J. Yang, "Time-domain acoustic contrast control design with response differential constraint in personal audio systems," *J. Acoust. Soc. Am.*, vol. 135, no. 6, pp. EL252–EL257, Jun. 2014.

[16] D. H. M. Schellekens, M. B. Møller, and M. Olsen, "Time domain acoustic contrast control implementation of sound zones for low-frequency input signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, Mar. 2016, pp. 365–369.

[17] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4518–4525, Jun. 2012.

[18] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "Time domain optimization of filters used in a loudspeaker array for personal audio," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 11, pp. 1869–1878, Nov. 2015.

[19] J. Benesty, M. G. Christensen, and J. R. Jensen, *Signal enhancement with variable span linear filters*. Springer, 2016, vol. 7.

[20] J. K. Nielsen, T. Lee, J. R. Jensen, and M. G. Christensen, "Sound zones as an optimal filtering problem," in *Proc. 52th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2018, pp. 1075–1079.

[21] T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, "A unified approach to generating sound zones using variable span linear filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Calgary, AB, Canada, Apr. 2018, pp. 491–495.

[22] J. R. Jensen, J. Benesty, and M. G. Christensen, "Noise reduction with optimal variable span linear filters," *IEEE/ACM Trans. Audio, Speech, and Lang. Process.*, vol. 24, no. 4, pp. 631–644, Apr. 2015.

[23] B. N. Parlett, *The Symmetric eigenvalue problem*. Society for Industrial and Applied Mathematics, Jan. 1998, vol. 20.

[24] S. J. Elliott and J. Cheer, "Regularisation and robustness of personal audio systems," ISVR Technical Memorandum 995, Tech. Rep., 2011.

[25] M. Poletti, "An investigation of 2-D multizone surround sound systems," in *Proc. 125th Conv. Audio Eng. Soc.*, San Francisco, CA, USA, Oct. 2008.

[26] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Carnegie Mellon University, Pittsburgh, PA, USA, Tech. Rep., Aug. 1994.

[27] K. E. Atkinson, *An introduction to numerical analysis*, 2nd ed. John Wiley & Sons, Jan. 1989.

[28] J. Francombe, P. Coleman, M. Olik, K. R. Baykaner, P. J. B. Jackson, R. Mason, S. Bech, M. Dewhirst, J. A. Pedersen, and M. Dewhirst, "Perceptually optimized loudspeaker selection for the creation of personal sound zones," in *Proc. 52nd Int. Conf. Audio. Eng. Soc.*, Guildford, UK, Sep. 2013.